

DOCUMENT RESUME

ED 320 930

TM 015 149

AUTHOR Huberty, Carl J.; And Others
TITLE The Study of Multivariate Group Contrasts.
PUB DATE Apr 90
NOTE 32p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Analysis of Variance; Comparative Analysis; Computer Software; *Computer Uses in Education; Construct Validity; *Data Analysis; *Groups; *Multivariate Analysis; *Research Methodology; Simulation
IDENTIFIERS *Contrast Analysis

ABSTRACT

What a multivariate contrast analysis should entail and how such an analysis should be conducted are discussed. The three questions encompassing a contrast analysis are: (1) whether the contrast effects are generalizable; (2) which variables contribute most to the contrast effects; and (3) what is the construct underlying the contrast effects. Three sets of results are needed for a contrast analysis using computer packages: test statistic and P value; outcome variable ordering; and underlying construct definition. Some helpful statistical package programs are discussed, including: BMDP, SAS, and SPSS-X-DISCRIMINANT. Two real data sets illustrate the use of multivariate contrast analyses. In one of these, a one-factor data set with three levels, group sizes were 35, 81, and 37. The other was a two-factor data set with four outcome variables and 18 observation vectors. Conducting a partial pairwise contrast analysis when the data do not satisfy the equal covariance matrices condition is discussed. Seven data tables are included. (SLD)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

The Study of
Multivariate Group Contrasts

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- ☒ This document has been reproduced as
received from the person or organization
originating it.
☐ Minor changes have been made to improve
reproduction quality.

- Points of view or opinions stated in this docu-
ment do not necessarily represent official
OERI position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

CARL HUBERTY

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Carl J Huberty

Tungshan F. Chou

Elisa B. Benitez

University of Georgia

325 Aderhold Hall
University of Georgia
Athens, GA 30602

BEST COPY AVAILABLE

Paper presented at the annual meeting of the American Educational
Research Association, Boston, April 1990.

ED320930

TM1015149

The Study of Multivariate Group Contrasts

Univariate Contrasts

The study of group contrasts in univariate contexts is very common. Nary a textbook that covers simple (i.e., one-way) analysis of variance (ANOVA) may be found that does not contain a discussion of group contrasts, some of which are called multiple comparisons. In virtually every such textbook, contrasts are covered after the omnibus ANOVA F test is covered. In fact, many textbook authors lead the reader to believe that contrasts are only studied after the one-way ANOVA F test indicates that significant group mean differences exist. Of course, there is no logical reason for conducting the omnibus test prior to the study of contrasts. [Some methodologists might argue that there are statistical reasons for preceding the study of contrasts with an omnibus test.] If there are specific research questions of interest to the researcher, and these may be addressed via the study of contrasts, then it seems reasonable to go directly to contrast analyses, and by-pass the omnibus test.

The notion of a univariate group contrast is quite familiar to nearly all empirical researchers. First of all, groups are being contrasted or compared in terms of the outcome variable group means. A sample contrast, C , then, may be defined as a particular linear combination of group means:

$$C = \sum_{i=1}^k a_i \bar{Y}_i,$$

where k is the total number of groups in the design, and the a_i are real numbers (usually integers) such that $\sum_{i=1}^k a_i = 0$. With $k=4$, three examples of contrasts are:

1. $\mu_2 = \mu_3$,
2. $2\mu_1 - \mu_2 - \mu_3$, and
3. $-3\mu_1 - \mu_2 + \mu_3 + 3\mu_4$

The first may be called a pairwise contrast, while the second and third are complex contrasts; the third is specifically a linear trend contrast. The constraint $\sum a_i = 0$ implies that all contrasts may be viewed as two-group comparisons. In contrast #2 above, for example, one may write $2\mu_1 - (\mu_2 + \mu_3)$; thus the comparison involves Group 1 on one hand and Groups 2 and 3 on the other hand. Therefore, with each contrast we may associate a hypothesis degree-of-freedom value of unity.

Many introductory textbooks present a number of methods that may be utilized in a contrast analysis -- popular methods are those associated with such names as Dunn, Fisher, Kramer, Newman-Keuls, Scheffe, and Tukey. As discussed elsewhere (Huberty & Morris, 1988), however, it may be argued that nothing more than a simple Student t statistic is needed in nearly every research situation. In evaluating the t statistic value, of course, one must take into consideration the total number of contrasts being tested.

Multivariate Contrasts

The notion of a multivariate group contrast is the same as that of a univariate group contrast except that vectors of outcome

variable means are involved. The multivariate counterpart of univariate contrast #2 above would be denoted as

$$2\mu_1 - \mu_2 - \mu_3.$$

The focus herein is on contrasting groups simultaneously across the multiple outcome variables, as opposed to contrasting groups on single outcome variables (as discussed by Ramsey, 1980). It may be argued that if univariate contrasts are those of primary interest, then no multivariate analyses would be called for (Huberty & Morris, 1989).

Multivariate contrasts are very seldom of concern, apparently, in publications that involve a multivariate analysis of variance (MANOVA). This is, perhaps, not too surprising since very few textbooks discuss multivariate contrasts in any great detail. Of the 15 applied multivariate textbooks collectively reviewed by Huberty and Barton (in press), only six even hint at multivariate group contrasts. Of these six, one (Dunteman, 1984, pp. 97-103) suggests only trend contrasts, and two (Harris, 1985, pp. 159-161; Marasciulo & Levin, 1983, pp. 324-329) discuss group contrasts of means on linear composites. Two textbooks discuss contrasts in non-specific terms (Cliff, 1987, pp. 402-405, 414-416; Lunneborg & Abbott, 1983, pp. 346-353), while another (Stevens, 1986, pp. 168-183) goes into some detail on conducting multivariate contrast statistical tests.

Three ideas mentioned in the opening section apply also to a multiple-outcome-variable context. One is that it is not necessary to conduct an omnibus mean vector comparison -- that is, a one way

MANOVA -- prior to the study of multivariate group contrasts. This idea, as in a univariate context, is very relevant in a group comparison situation -- a general omnibus group comparison question is only seldom of interest. More interesting questions are usually more specific, those reflected as contrasts. The second idea applicable in a multivariate context is that of employing a single statistic in all situations -- the test statistic may be, equivalently, the Hotelling two-group T^2 , the Wilks lambda, or the Mahalanobis D^2 . The error covariance matrix typically used with any of these statistics is the pooled covariance matrix; separate covariance matrices may, however, be used. The third idea involves multiple contrast testing; if m tests are conducted, then the tail probability value for each test is to be multiplied by m .

With a multivariate situation, however, it is not sufficient to merely conduct the statistical test for a given contrast. Two additional sets of results may very well be informative. In particular, a researcher may be interested in variable ordering and/or in construct definition. The variable ordering problem pertains to determining the relative importance or contribution of the outcome variables. More than one interpretation of variable relative contribution has been proffered (Huberty & Wisenbaker, 1990). The interpretation considered here is relative contribution to grouping variable effects; i.e., to the variability in the outcome variable vector attributable to the grouping variable, as reflected in the complement of the Wilks lambda. For a contrast, this means relative contribution to the contrast statistical

effect; or, equivalently, to the contrast Wilks lambda value. Thus, for p outcome variables, one could conduct p contrast analyses each with $p-1$ variables. The analysis that yields the least separation, the largest Wilks lambda, indicates that the deleted variable is the most important one. As will be seen shortly, however, only one analysis needs to be done.

The construct definition problem pertains to identifying/labeling/naming the construct that underlies the contrast effect. In a contrast (i.e., two-group) analysis, there is one linear discriminant function (LDF). It is this linear composite of the outcome variables that is a basis for the construct. The statistic useful for construct definition is the structure correlation -- the correlation between scores on an outcome variable and scores on the LDF. These within-groups correlations are sometimes called loadings.

The purposes of this paper are to present an interpretation of what a contrast analysis should entail, and to discuss in detail how a multivariate contrast analysis can be conducted. There are three questions encompassing a "contrast analysis":

- 1) Are the contrast effects generalizable? In the context of a pairwise contrast, the question is: Are the two mean vectors significantly different?
- 2) Which variables contribute most to the contrast effects?
- 3) What is the construct underlying the contrast effects? That is, with what construct do we associate the group mean vector differences?

A discussion is presented on how to obtain the three sets of results for a contrast analysis via computer packages: (1) test statistic and P value; (2) outcome variable ordering; and (3) underlying construct definition. Real data sets are utilized.

Some Helpful Statistical Package Programs

The three statistical computer packages considered here are BMDP, SAS, and SPSS-X. As discussed in the last section, three sets of results are desirable to complete a contrast analysis: 1) statistical test; 2) variable ordering; and 3) construct definition. An indication of which program in each package that provides which type of result is given in Table 1. The information obtainable from each of the four programs will be illustrated via some exemplary analyses.

Insert Table 1 About Here

A special note needs to be made about the SPSS-X DISCRIMINANT program. To use this program for contrast analyses, two "runs" need to be made -- see Appendix A. The first run is used to obtain some matrix information: a pooled (or, error) correlation matrix, group sizes, and group means and standard deviations for all outcome variables. This matrix information, then, may be used in the second run to obtain some contrast analysis results. There are some problems with this direct use of the second run. One is that the resulting Wilks lambda and, necessarily, the F statistic values may be incorrect; similarly for the resulting F-to-remove values.

[The reason for this is that the computational algorithm transforms the error correlation matrix to an error SSCP matrix using the individual group sizes (for only those two groups involved in the contrast) rather than the total group size.] The structure r values obtained with the direct use of the two runs are, however, correct. The correct test and ordering information may be obtained by editing the output matrix file so that the sizes of the groups involved are adjusted to obtain a correct multiplier to transform the correlation matrix.

The use of the two runs for conducting a complex contrast analysis requires additional editing of the matrix output file. This editing involves the construction of two "groups" by combining (according to the specified contrast) original group n 's, means, and standard deviations. Such editing is not detailed herein.

Examples

Two real data sets will be utilized to illustrate the conduct of multivariate contrast analyses. For each data set it is initially assumed that the population covariance matrices are equal, and that there are no outlying or influential observation vectors. It should be pointed out that the data sets are currently being used for methodological and not substantive reasons. Therefore, outcome variables will simply be identified by numbers. For each data set, two types of contrasts will be considered, a pairwise contrast and a complex contrast.

Data Set One

This is a one-factor data set with three levels, and with 13 outcome variables. Group sizes are 35, 81, and 37. The three population covariance matrices are assumed to be equal -- the approximate F value ($df = 182/29061$) for the Box test is 1.15, $P = .077$. Two contrasts considered are $\mu_1 - \mu_2$ and $\mu_1 + \mu_2 - 2\mu_3$. Partial set-up command statements for the four package programs considered here are given in Appendix A. A couple comments regarding some of the commands are called for. First, the low ENTER and REMOVE values for BMDP 7M and FIN and FOUT values for SPSS-X DISCRIMINANT are used merely to virtually ensure that all variables are entered into the analysis. [Alternatively with BMDP 7M, the statement $FORCE = 2$ may be used in the DISCRIMINANT paragraph to ensure that all variables will be entered.] Second, there are some seven ways with SPSS-X MANOVA to specify a contrast, and SPECIAL is one which calls for a set of three (the number of groups) 1's, followed by the actual contrast of interest, followed by a third (because there are three groups) set of weights; this "extra" set of weights is arbitrary except that they cannot represent a contrast that is a linear composite of earlier contrasts (with four groups, two "extra" contrasts are needed). It should be noted that analyses of the two contrasts specified above may be accomplished with a single SPSS-X MANOVA contrast command, e.g.,

```
CONTRAST (GRP) = SPECIAL (1 1 1 1 -1 0 1 1 -2)
```

Both contrast analyses may also be accomplished with a single BMDP

7M set of commands by simply using two CONTRAST commands in the DISCRIMINANT paragraph; similarly for SAS GLM. [Additional descriptive information may be obtained using a variety of commands with each of the four programs. Some descriptive information for Data Set One is presented in Table 2.]

Insert Table 2 About Here

$\mu_1 - \mu_2$. The Wilks lambda value for this contrast is common for BMDP 7M, SAS GLM, and SPSS-X MANOVA (viz, .42827), as is the approximate F value (viz, $F(13, 138) = 14.171$); the P value is given by the two latter programs (to three places it is .000), and eta-squared = $1 - .42827 = .572$. It seems reasonable to conclude that the contrast effects are generalizable.

Our experience with the four programs tells us that an easy way to obtain correct variable ordering information is to use BMDP 7M. The preferred index of variable ordering (see Huberty & Wisenbaker, 1990) is the F TO REMOVE -- the F value for the i th outcome variable, $F_{(i)}$, reflects that group separation yielded by all outcome variables less Y_i . These F values are reported in the last "step" of the analysis. For this data set, the 13 F-to-remove values are given in Table 3. It is perhaps clear from Table 3

Insert Table 3 About Here

that V12 contributes most to Group 1 versus Group 2 mean differences, while V11 is the second most important variable, followed by V4 and V5, with V1, V6, and V8 being least important. See Huberty and Wisenbaker (1990) for a more precise method of rank-ordering variables using F-to-remove values.

The structure r values reported by the two SPSS-X programs are identical (except for algebraic sign) and are given in Table 3. From these structure r 's one may conclude that the construct underlying the Group 1 - Group 2 contrast effects is basically defined by V12 along with, possibly, V5 and V11. [That these three variables are also some of the more important ones -- as determined by F-to-remove values -- is not a result that one should expect with all data sets.]

$\mu_1 + \mu_2 - 2\mu_3$. The Wilks lambda value for this contrast is common for BMDP 7M, SAS GLM, and SPSS-X MANOVA (viz, .31573), as is the approximate F value (viz, $F(13,138) = 23.006$); the P value given by the two latter programs is .000, and eta squared $\hat{=}$.685. Again, it seems reasonable to conclude that the contrast effects are generalizable. [Assuming the two contrasts discussed are the only ones to be considered, the two P values should technically be multiplied by 2; for the two tests conducted here, statistical conclusions would not change.]

The F-to-remove values for this contrast -- obtainable from BMDP 7M -- are reported in Table 4. For this contrast, V12

Insert Table 4 About Here

is most important, followed by V11, then by V4, with V1, V3, and V6, V8, and V10 contributing least to the contrast effects.

The structure r values reported by the two SPSS-X programs are also given in Table 4. From these loadings we see that the construct underlying the contrast effects is essentially defined by V11 and V12 along with, but to a lesser extent, V5.

Data Set Two

This is a two-factor (A with two levels, B with four levels) data set with four outcome variables and 18 observation vectors in each of the eight cells. The Box test for the equality of the eight population covariance matrices yielded an approximate $F(70, 25322) = 2.516$ with $P = .000$. [In the face of this result, we still proceed with a conventional analysis for illustrative purposes.] The data support the nonexistence of a multivariate $A \times B$ interaction -- $\lambda = .9296$, $F(12, 352) = 0.82$, $P = .629$. Therefore, contrasts among levels of B were studied in terms of the marginal mean vectors. The two contrasts considered are $\mu_{a1} - \mu_{a3}$ and $3\mu_{a1} - \mu_{a2} - \mu_{a3} - \mu_{a4}$. Partial set-up command statements for the three package programs considered here are given in Appendix B. Just as for Data Set One, the two BMDP 7M CONTRAST commands may be used in one DISCRIMINANT paragraph; similarly with SAS GLM. This very idea was used with SPSS-X MANOVA for the two contrasts. Note that in the SAS GLM MODEL statement, only a single "factor" (namely, AB) is used. This was done since in the DATA step, levels of AB were defined via levels of A and levels of B:

```

If A = 1 AND B = 1 THEN AB = 1;
If A = 1 AND B = 2 THEN AB = 2;
.
.
.
If A = 2 AND B = 4 THEN AB = 8;

```

Recall that with MANOVA the user needs to fill in the CONTRAST command -- we used SPECIAL here -- with as many sets of weights, beyond the set of 1's, as there are degrees of freedom for the overall "model" (we used orthogonal polynomial coefficients).

Some descriptive information for Data Set Two is presented in Table 5.

Insert Table 5 About Here

$\mu_{11} - \mu_{22}$. The Wilks lambda value for this contrast obtained from all three programs is .92327 (eta-squared = .077) with $F(4, 133) = 2.763$ and $P \hat{=} .030$.

The F-to-remove (from BMDP 7M) and structure r (from SPSS-X MANOVA) values are given in Table 6. Clearly, V1 contributes most

Insert Table 6 About Here

to these contrast effects, and V3 is next most important, with V4 contributing least. The construct underlying Groups 1 and 3 mean differences is defined by V1 and V3.

$3\mu_{11} - \mu_{22} - \mu_{33} - \mu_{44}$. The Wilks lambda value for this contrast is .88892 (eta squared = .111) with $F(4, 133) = 4.155$ and $P \hat{=} .003$. [Note that considering the fact that two contrasts were

tested, each of the two P values would be multiplied by 2. Thus, the effects for the pairwise contrast may not be considered generalizable.]

The F-to-remove and structure r values are given in Table 7.

Insert Table 7 About Here

Clearly V1 is most important in the sense of defining the construct underlying the effects. [Whereas V3 contributed to mean differences between B1 and B2 and to the definition of the underlying construct, either contribution by V3 with respect to the complex contrast is relatively little.]

Unequal Covariance Matrices

It is generally recognized that the Box test for the equality of the population covariance matrices is very powerful -- witness the second degrees-of-freedom values for the two F test results presented earlier. So, it seems seldom that real data based on anything but a very small design ("few" outcome variables, groups, and cases) will satisfy this condition. [Was Data Set One a "surprise"?] Thus, researchers may find themselves confronted with contrast analyses for data with unequal covariance matrices.

What is intended here is to demonstrate how a (partial) pairwise contrast analysis may be conducted when the data do not satisfy the equal covariance matrices condition. What will be employed is a two-sample Hotelling T^2 statistic when the condition is not met. The statistic is:

$$T^2 = (\bar{Y}_1 - \bar{Y}_2)' (S_1/n_1 + S_2/n_2)^{-1} (\bar{Y}_1 - \bar{Y}_2),$$

where $\bar{Y}_j$, S_j , and n_j denote the outcome variable mean vector, covariance matrix, and size of Group j , respectively. An SAS program to compute a T^2 value is given in Appendix C. The T^2 value for the contrast $\mu_1 - \mu_2$ with Data Set One is 200.243. Algina and Tang (1988) studied two different approximate referent distributions for T^2 , while Johnson and Wichern (1988, p. 229) suggest, simply, a chi-squared distribution with p (the number of outcome variables) degrees of freedom. Although we concur that an F distribution should, in practice, be utilized, for our purposes here we will use the chi-squared distribution (with $df=13$). The P value is .000, indicating generalizable contrast effects. This comprises the obtainable statistical test information.

But, what about information pertaining to variable ordering and to construct definition? We know of no proposals for defining constructs when covariance matrices are not equal. And it is inappropriate to rely on F -to-remove values for ordering with unequal covariance matrices. What can be done, however, is to conduct p T^2 tests, each with $p-1$ outcome variables. Let $T_{(i)}$ denote the value of T^2 with Y_i deleted. The outcome variable for which $T_{(i)}$ is least would be judged to be the most important variable.

Of course, a T^2 value is obtainable for any kind of contrast pairwise or complex. Also, P value adjustments should be implemented for multiple tests.

Comments

If a research worker is interested in the effect of some grouping variable(s) on each outcome variable in a collection of conceptually independent outcome variables, then a series of univariate analyses would be called for. On the other hand, if some system of outcome variables is being studied, then it seems reasonable to conduct a multivariate analysis. In such a MANOVA context, there may be research situations where omnibus questions are of some interest. That is, interest may be on main effects questions -- this may be in addition to interaction effects in a two- or higher-way layout. If so, it would be helpful to obtain more information than merely that pertaining to a statistical test. What is being referred to is information on relative variable contribution to resulting main effects (or, interaction effects, for that matter), and on any latent construct(s) that underlie the effects. This information would be helpful whether one were researching in an exploratory mode or in a confirmatory mode.

If a researcher's design involved multiple levels of a grouping variable, then perhaps a very reasonable analysis approach would involve the study of group contrasts, still in a multivariate setting. Of course, if only contrast effects are of interest (and not omnibus effects), then the researcher should not feel compelled to conduct omnibus analysis first. In such a situation, he/she should proceed directly to contrast analyses, the focus of this paper. Just as for omnibus analyses and so emphasized in this paper, the examination of three types of results is strongly

recommended: statistical test, variable ordering, and construct definition.

It is disappointing that no single statistical computer program package may be utilized directly and readily to obtain all three types of results for a contrast analysis. It is the case, however, if a researcher is adept at editing data output files, the SPSS-X DISCRIMINANT is a possibility. Editing adeptness is particularly needed for the complete study of complex contrasts using SPSS-X DISCRIMINANT. Obtaining the first two types of information (statistical test and variable ordering) when dealing with unequal covariance matrices requires some programming skills.

There may be situations in which a researcher has included multiple systems of outcome variables. For example, he/she may have some interest variables, some locus of control variables, some achievement variables, and/or some motivation variables. In such a situation multiple multivariate analyses -- to study omnibus and/or contrast effects -- would very well be appropriate. Jensen (1972) presents some technical details for such analyses.

References

- Algina, J., & Tang, K. L. (1988). Type I error rates for Yao's and James' tests of equality of mean vectors under variance-covariance heteroscedasticity. Journal of Educational Statistics, 13, 281-290.
- Cliff, N. (1987). Analyzing multivariate data. New York: Harcourt Brace Jovanovich.
- Dunteman, G. H. (1984). Introduction to multivariate analysis. Beverly Hills, CA: Sage.
- Harris, R. J. (1985). A primer of multivariate statistics. Orlando, FL: Academic Press.
- Huberty, C. J., & Barton, R. M. (in press). Applied multivariate statistics textbooks. Applied Psychological Measurement.
- Huberty, C. J., & Morris, J. D. (1988). A single contrast test procedure. Educational and Psychological Measurement, 48, 567-578.
- Huberty, C. J., & Morris, J. D. (1989). Multivariate Analysis Versus Multiple Univariate Analyses. Psychological Bulletin, 105, 302-308.
- Huberty, C. J., & Wisenbaker, J. M. (1990, April). Variable importance in MANOVA and DA. Paper presented at the annual meeting of the American Educational Research Associate, Boston.
- Jensen, D. R. (1972). Some simultaneous multivariate procedures using Hotelling's T^2 statistics. Biometrics, 28, 39-53.

- Lunneborg, C. E., & Abbott, R. D. (1983). Elementary multivariate analysis for the behavioral sciences. New York: North-Holland.
- Marascuilo, L. A., & Levin, J. R. (1983). Multivariate statistics in the social sciences. Monterey, CA: Brooks/Cole.
- Ramsey, P. H. (1980). Choosing the most powerful pairwise multiple comparison procedure in multivariate analysis of variance. Journal of Applied Psychology, 65, 317-326.
- Stevens, J. (1986). Applied multivariate statistics for the social sciences. Hillsdale, NJ: Erlbaum.

Appendix A

Some Program Commands for Data Set One

$$\mu_1 - \mu_2$$

$$\mu_1 + \mu_2 - 2\mu_3$$

BMDP 7M

DISCRIMINANT

ENTER = .00001.
REMOVE = .00001.
CONTRAST = 1, -1, 0.

DISCRIMINANT

ENTER = .00001.
REMOVE = .00001.
CONTRAST = 1, 1, -2.

SAS GLM

PROC GLM; CLASS GRP;

(Model statement here)
MANOVA H = GRP;
CONTRAST '1 VS 2' GRP 1 -1 0;

PROC GLM; CLASS GRP;

(Model statement here)
MANOVA H = GRP;
CONTRAST '1 2 VS 3' GRP 1 1 -2

SPSS-X
MANOVA

MANOVA (Variable list here)

CONTRAST (GRP) =
SPECIAL (1 1 1 1 -1 0 1 0 -1)/
PARTITION (GRP)/
DISCRIM = COR/
DESIGN = GRP (1), GRP (2)/

MANOVA (Variable list here)

CONTRAST (GRP) =
SPECIAL (1 1 1 1 1 -2 1 -1 0)/
PARTITION (GRP)/
DISCRIM = COR/
DESIGN = GRP (1), GRP (2)/

SPSS-X
DISCRIMINANT

DISCRIMINANT GROUPS = GRP (1,3)/

VARIABLES (List here)
METHOD = WILKS/FIN = .00001/
FOUT = .00001/
MATRIX = OUT (*)

(See the text)

DISCRIMINANT GROUPS = GRP (1,2)/

VARIABLES (List here)
METHOD = WILKS/FIN = .00001/
FOUT = .00001/
MATRIX = IN (*)

Appendix B

Some Program Commands for Data Set Two

	$\mu_{B1} - \mu_{B3}$		$3\mu_{B1} - \mu_{B2} - \mu_{B3} - \mu_{B4}$
BMDP 7M	VARIABLE NAMES ARE GRP, X1,X2,X3,X4. GROUPING IS GRP./ GROUP CODES(1) ARE 1 TO 8. NAMES(1) ARE G1,G2,G3,G4,G5, G6,G7,G8./ DISCRIMINANT ENTER = .00001 REMOVE = .00001 CONTRAST = 1,0,-1,0,1,0,-1,0		VARIABLE NAMES ARE GRP, X1,X2,x3,x4. GROUPING IS GRP./ GROUP CODES(1) ARE 1 TO 8. NAMES(1) ARE G1,G2,G3,G4,G5, G6,G7,G8./ DISCRIMINANT ENTER = .00001 REMOVE = .00001 CONTRAST = 3,-1,-1,-1,3,-1,-1,-1.
SAS GLM	PROC GLM; CLASS AB; MODEL X1 X2 X3 X4 = AB; MANOVA H = AB; CONTRAST '1 VS 3' AB 1 0 -1 0 1 0 -1 0;		PROC GLM; CLASS AB; MODEL X1 X2 X3 X4 = AB; MANOVA H = AB; CONTRAST '1 VS 2 3 4' AB 3 -1 -1 -1 3 -1 -1 -1;
SPSS-X MANOVA	MANOVA (Variable list here) CONTRAST (GRP) = SPECIAL (1 1 1 1 1 1 1 1 1 0 -1 0 1 0 -1 0 3 -1 -1 -1 3 -1 -1 -1 -7 -5 -3 -1 1 3 5 7 7 1 -3 -5 -5 -3 1 7 -7 5 7 3 -3 -7 -5 7 7 -13 -3 9 9 -3 -13 7 -7 23 -17 -15 15 17 -23 7)/ PARTITION (GRP)/ DISCRIM = COR/ DESIGN = GRP(1), GRP(2), GRP(3), GRP(4), GRP(5), GRP(6), GRP(7)/		(See left side)

Appendix C

SAS Program for Hotelling T²

```

PROC SORT DATA = TWO; BY GRP;
PROC CORR COV DATA = TWO OUTP = SCOV NOPRINT NOSIMPLE; BY GRP;
  VAR (Variable names here);
DATA B1; SET SCOV;
  IF GRP = 1 THEN; ELSE DELETE;
DATA B2; SET SCOV;
  IF GRP = 2 THEN; ELSE DELETE;

PROC IML; RESET;
  USE B1;
  NAME = [Variable names here];
  READ ALL WHERE (___ TYPE ___ = 'COV') INTO COV1 VAR NAME;
  USE B2;
  READ ALL WHERE (___ TYPE ___ = 'COV') INTO COV2 VAR NAME;

  M1 = [Group 1 means here];
  M2 = [Group 2 means here];

  T2 = (M1 - M2) * INV (COV1/35 + COV2/81) * (M1 - M2)';

PRINT, T2;

```

Table 1

Contrast Results Generated by the Statistical Packages

<u>Package Programs</u>	<u>Test</u>	<u>Result</u>	
		<u>Ordering</u>	<u>Construct</u>
BMDP 7M (1987)	Yes	Yes	No
SAS GLM (Version 5.18)	Yes	No	No
SPSS-X (Release 3.0)			
MANOVA	Yes	No	Yes
DISCRIMINANT	*	*	Yes

*See the text.

Table 2

Group Sizes, Means, Standard Deviations^a, and Error Correlations for Data Set One

	Group 1	Group 2	Group 3	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13
V1	2.90543 (.5045)	3.28741 (.5052)	3.33108 (.4906)	.4479	.5973	.5504	-.2900	.5596	.3717	.1486	.2648	.1811	-.2129	.0479	-.0488
V2	2.52000 (.6986)	2.85617 (.7989)	2.68676 (.7971)		.4259	.6741	-.2107	.5527	.2756	.5317	.2647	.3240	-.2062	-.0544	.1531
V3	2.89114 (.4768)	3.34321 (.5193)	3.20919 (.5548)			.5913	-.1988	.5062	.0517	.1802	.2229	.1188	-.2400	-.1054	-.0157
V4	2.65343 (.5431)	3.07395 (.5954)	2.95432 (.6895)				-.3360	.5107	.2457	.3695	.3196	.2906	-.3454	-.1514	-.0329
V5	3.17143 (.9848)	4.23457 (.7628)	4.97297 (1.1177)					-.3544	-.2647	-.1820	.0216	-.0355	.2932	.2609	.0268
V6	2.67629 (.6901)	3.20420 (.6389)	3.14757 (.6654)						.3898	.3189	.2398	.1961	-.0671	.0152	-.1650
V7	22.42857 (2.0904)	23.80247 (2.8610)	25.02703 (2.5109)							.3600	.3319	.3981	.0732	.0531	.0356
V8	23.65714 (5.7340)	24.79012 (4.9918)	24.27027 (5.3471)								.3774	.5251	-.0343	-.0731	.2420
V9	25.14286 (3.9567)	26.59259 (3.1416)	27.32432 (3.6746)									.5878	.1279	.1414	.0591
V10	23.31429 (4.1570)	24.62963 (4.5674)	26.05405 (4.6064)										.2012	.0635	.1521
V11	14.34286 (4.7832)	19.27160 (3.6229)	25.27027 (3.9627)											.2112	-.1582
V12	40.85714 (7.5504)	52.87654 (7.2170)	63.75676 (4.0098)												-.3459
V13	2.00000 (2.1557)	1.39506 (1.6022)	.48649 (.9894)												
n	35	81	37												

^aStandard deviations in parentheses.

Table 3

F-to-Remove and Structure r Values for $\mu_1 - \mu_2$ with Data Set One

<u>Variable</u>	<u>F-to-Remove</u>	<u>Structure r</u>
V1	.00	.266
V2	2.18	.151
V3	3.28	.304
V4	11.87	.242
V5	10.12	.408
V6	.97	.281
V7	1.59	.183
V8	.39	.075
V9	2.80	.146
V10	1.66	.102
V11	15.88	.431
V12	29.10	.629
V13	3.06	-.130

Table 4

F-to-Remove and Structure 4 Values for $\mu_1 + \mu_2 - 2\mu_3$ with Data Set One

<u>Variable</u>	<u>F-to-Remove</u>	<u>Structure r</u>
V1	.17	.134
V2	1.60	-.000
V3	.78	.051
V4	10.49	.043
V5	4.69	.401
V6	.09	.091
V7	2.65	.210
V8	.22	.003
V9	3.26	.121
V10	.89	.133
V11	32.60	.609
V12	43.13	.727
V13	1.37	-.213

Table 5

Means, Standard Deviations*, and Error Correlations for Data Set Two

	A1B2	A1B2	A1B3	A1B4	A2B1	A2B2	A2B3	A2B4	V2	V3	V4
V1	21.66667 (9.6162)	15.05556 (4.8806)	15.77778 (7.1419)	15.38889 (4.3540)	18.55556 (8.2549)	13.27778 (6.4334)	17.11111 (4.9096)	15.55556 (6.3914)	.2507	-.0976	.1589
V2	4.94444 (3.4209)	5.00000 (2.8284)	4.94444 (2.8589)	5.55556 (2.9946)	3.44444 (3.0912)	2.33333 (2.4734)	3.44444 (2.7272)	4.50000 (2.8336)		.1983	.0116
V3	4.88889 (7.9918)	4.44444 (9.8173)	1.50000 (1.8865)	3.88889 (4.4969)	3.66667 (4.9468)	2.50000 (3.2222)	1.55556 (1.6169)	6.22222 (8.8753)			.0969
V4	1.05556 (2.1275)	1.27778 (2.5392)	1.66667 (2.8492)	1.88889 (3.0657)	3.88889 (6.5159)	2.44444 (4.2596)	1.44444 (3.1849)	3.66667 (5.5836)			

Note. All cell sizes are 18.

*Standard deviations in parentheses.

Table 6

F-to-Remove and Structure r Values for $\mu_{B1} - \mu_{B3}$ with Data Set Two

<u>Variable</u>	<u>F-to-Remove</u>	<u>Structure r</u>
V1	6.56	-.689
V2	1.11	.000
V3	4.90	-.566
V4	.08	-.286

Table 7

F-to-Remove and Structure r Values for $3\mu_{B1} - \mu_{B2} - \mu_{B3} - \mu_{B4}$ with Data Set Two

<u>Variable</u>	<u>F-to-Remove</u>	<u>Structure r</u>
V1	15.60	-.892
V2	1.87	.044
V3	1.92	-.190
V4	.07	-.127
